HOMEWORK 2: SOLUTIONS

1.

Consider three events A_+, A_-, A_0 in the same probability space Ω with a probability P on it. Suppose that A_+ and A_- are conditionally independent given A_0. Show that $P(A_+|A_0 \cap A_-) = P(A_+|A_0)$.

Solution.

By definition, A_+ and A_- are conditionally independent given A_0, if

$$P(A_+ \cap A_-|A_0) = P(A_+|A_0)\ P(A_-|A_0)$$

We thus have

$$\begin{aligned} P(A_+|A_0 \cap A_-) = \frac{P(A_+ \cap A_0 \cap A_-)}{P(A_0 \cap A_-)} &= \frac{P(A_+ \cap A_0 \cap A_-)/P(A_0)}{P(A_0 \cap A_-)/P(A_0)} \\ &= \frac{P(A_+ \cap A_-|A_0)}{P(A_-|A_0)} \\ &= P(A_+|A_0). \end{aligned}$$

2.

Suppose that the sequence $X_0, X_1, X_2, \ldots$ of random variables with values in $\mathbb{Z}$ have the Markov property.

(i) Show that the sequence $X_0^3, X_1^3, X_2^3, \ldots$ also has the Markov property.

(ii) Given an example where the sequence $X_0^2, X_1^2, X_2^2, \ldots$ does not have the Markov property.

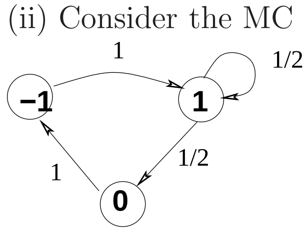
Hint: If $x^3 = y$ then $x = y^{1/3}$. For example, $x^3 = -8$ means that $x = -2$. But if $x^2 = y$ then x equals $\sqrt{|y|}$ or $-\sqrt{|y|}$. For example, $x^2 = 100$ means that $x = 10$ or $x = -10$.

Solution.

(i) We have

$$\begin{aligned} &P(X_{n+1}^3 = j|X_n^3 = i, X_{n-1}^3 = i_1, \ldots X_0^3 = i_0) \\ &\qquad = P(X_{n+1} = j^{1/3}|X_n = i^{1/3}, X_{n-1} = i_1^{1/3}, \ldots, X_0 = i_0^{1/3}) \\ &\qquad = P(X_{n+1} = j^{1/3}|X_n = i^{1/3}) \\ &\qquad = P(X_{n+1}^3 = j|X_n^3 = i) \end{aligned}$$

(ii) Consider the MC $(X_0, X_1, X_2, \ldots)$ with transition diagram:



Let us assume that

$$P(X_0 = 0) = P(X_0 = -1) = 1/2.$$

We have

$$P(X_2^2 = 1 | X_1^2 = 1, X_0^2 = 0) = P(X_{n+1}^2 = 1 | X_n = 1, X_{n-1} = 0) = 1.$$

On the other hand,

$$P(X_2^2 = 1 | X_1^2 = 1) = 3/4.$$

Since the two are not equal, the process is not Markovian.

To see the computation of the last probability, do this:

$$P(X_2^2 = 1 | X_1^2 = 1) = \frac{P(X_2^2 = 1, X_1^2 = 1)}{P(X_1^2 = 1)}$$

Now,

$$\begin{aligned} P(X_2^2 = 1, X_1^2 = 1) &= \frac{1}{2} P(X_2^2 = 1, X_1^2 = 1 | X_0 = 0) + \frac{1}{2} P(X_2^2 = 1, X_1^2 = 1 | X_0 = -1) \\ &= \frac{1}{2} P(X_2 = 1, X_1 = -1 | X_0 = 0) + \frac{1}{2} P(X_2 = 1, X_1 = -1 | X_0 = -1) \\ &= \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot \frac{1}{2} = \frac{3}{4}. \end{aligned}$$

On the other hand, if $X_0 = 0$ then $X_1 = 1$. If $X_0 = 1$ then $X_1 = \pm 1$. In either case $X_1^2 = 1$ for sure. So $P(X_1^2 = 1) = 1$.

3. ---

In the Dark Ages, Harvard, Dartmouth, and Yale admitted only male students. Assume that, at that time, 80 percent of the sons of Harvard men went to Harvard and the rest went to Yale, 40 percent of the sons of Yale men went to Yale, and the rest split evenly between Harvard and Dartmouth; and of the sons of Dartmouth men, 70 percent went to Dartmouth, 20 percent to Harvard, and 10 percent to Yale.

(i) Find the probability that the grandson of a man from Harvard went to Harvard.

(ii) Modify the above by assuming that the son of a Harvard man always went to Harvard. Again, find the probability that the grandson of a man from Harvard went to Harvard.

(iii) Find the stationary distribution(s) of the Markov chain.

Solution. (i) The probability that the grandson of a man from Harvard went to Harvard equals

$$0.8^2 + 0.2 \times 0.3 = 0.7.$$

(ii) Now the probability that the grandson of a man from Harvard went to Harvard equals 1.

(iii) We let X_n ($n = 0, 1, 2, \ldots$) be the university attended by the the n-th generation son. This is, by assumption, a Markov chain with values in $S = \{H, D, Y\}$.

In case (i) the transition probability matrix (with this order of states (H, D, Y)) is

$$\mathsf{P} = \begin{pmatrix} 0.8 & 0 & 0.2 \\ 0.2 & 0.7 & 0.1 \\ 0.3 & 0.3 & 0.4 \end{pmatrix}$$

Any stationary distribution π satisfies $\pi\mathsf{P} = \pi$, i.e.

$$\begin{aligned} \pi(H) &= 0.8\pi(H) + 0.2\pi(D) + 0.3\pi(Y) \\ \pi(D) &= 0.7\pi(D) + 0.3\pi(Y) \\ \pi(Y) &= 0.2\pi(H) + 0.1\pi(D) + 0.4\pi(Y) \end{aligned}$$

These equations are linearly dependent, so I can always omit one of them: I choose to omit the last one. The second equation gives $\pi(D) = \pi(Y)$. The first one gives $0.2\pi(H) = (0.2 + 0.3)\pi(D) = 0.5\pi(D)$, so $\pi(H) = (5/2)\pi(D)$. In addition,

$$1 = \pi(H) + \pi(D) + \pi(Y) = (5/2 + 1 + 1)\pi(D) = 9/2\pi(D)$$

whence $\pi(D) = 2/9$. So the stationary distribution $\pi = (\pi(H), \pi(D), \pi(Y))$ is given by

$$\pi(H) = 5/9, \quad \pi(D) = 2/9, \quad \pi(Y) = 2/9.$$

In case (ii) the transition probability matrix is

$$\mathsf{P} = \begin{pmatrix} 1 & 0 & 0 \\ 0.2 & 0.7 & 0.1 \\ 0.3 & 0.3 & 0.4 \end{pmatrix}.$$

Any stationary distribution π satisfies $\pi\mathsf{P} = \pi$, i.e.

$$\begin{aligned} \pi(H) &= \pi(H) + 0.2\pi(D) + 0.3\pi(Y) \\ \pi(D) &= 0.7\pi(D) + 0.3\pi(Y) \\ \pi(Y) &= 0.1\pi(D) + 0.4\pi(Y) \end{aligned}$$

These equations are linearly dependent, so I can always omit one of them: I choose to omit the last one. But observe that $\pi(H)$ cancels from the first one, which means that it can be chosen arbitrarily: set $\pi(H) = p$, where p is any number, such that $0 \le p \le 1$. The second one gives $\pi(D) = \pi(Y)$. In addition,

$$1 = \pi(H) + \pi(D) + \pi(Y) = p + 2\pi(D),$$

whence $\pi(D) = (1 - p)/2$. We then have

$$\pi(H) = p, \quad \pi(D) = (1 - p)/2, \quad \pi(Y) = (1 - p)/2.$$

Since p is arbitrary, we have an infinite number of stationary distributions.

4.

Consider an experiment of mating rabbits. We watch the evolution of a particular gene that appears in two types, G or g. A rabbit has a pair of genes, either GG (dominant), Gg (hybrid–the order is irrelevant, so gG is the same as Gg) or gg (recessive). In mating two rabbits, the offspring inherits a gene from each of its parents with equal probability. Thus, if we mate a dominant (GG) with a hybrid (Gg), the offspring is dominant with probability 1/2 or hybrid with probability 1/2.
Start with a rabbit of given character (GG, Gg, or gg) and mate it with a hybrid. The offspring produced is again mated with a hybrid, and the process is repeated through a number of generations, always mating with a hybrid.
(i) Write down the transition probabilities of the Markov chain thus defined.
(ii) Assume that we start with a hybrid rabbit. Let μ_n be the probability distribution of the character of the rabbit of the n-th generation. In other words, $\mu_n(GG), \mu_n(Gg), \mu_n(gg)$ are the probabilities that the n-th generation rabbit is GG, Gg, or gg, respectively. Compute μ_1, μ_2, μ_3. Can you do the same for μ_n for general n?
(iii) Find the stationary distribution(s) of the Markov chain.

Solution. (i) We consider a Markov chain with values in $S = \{GG = 1, Gg = 2, gg = 3\}$ and transition probability matrix

$$\mathsf{P} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

(ii) We start with X_0 being distributed as $\mu_0 = (\mu_0(1), \mu_0(2), \mu_0(3)) = (0, 1, 0)$, and, letting $mu_n = (\mu_0(n), \mu_0(n), \mu_0(n))$ be the distribution of X_n, we have

$$\mu_n = \mu_0 \mathsf{P}^n.$$

So

$$\mu_1 = \mu_0 \mathsf{P} = (0, 1, 0) \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} = (\tfrac{1}{4}, \tfrac{1}{2}, \tfrac{1}{4}).$$

$$\mu_2 = \mu_0 \mathsf{P}^2 = (\mu_0 \mathsf{P})\mathsf{P} = \mu_1 \mathsf{P} = (\tfrac{1}{4}, \tfrac{1}{2}, \tfrac{1}{4}) \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} = (\tfrac{1}{4}, \tfrac{1}{2}, \tfrac{1}{4}).$$

So we see that $\mu_2 = \mu_1$. Therefore

$$\mu_3 = \mu_2 \mathsf{P} = \mu_1 \mathsf{P} = \mu_2 = \mu_1.$$

And so, for all $n \geq 1$,

$$\mu_n = \mu_1 = (\tfrac{1}{4}, \tfrac{1}{2}, \tfrac{1}{4}).$$

(iii) Since $\mu_2 = \mu_1$ we have

$$\mu_1 = \mu_1 \mathsf{P}$$

and so μ_1 is a stationary distribution. We can see that there are no other stationary distributions by seeing, for example, that the nullity of the system of equations $\pi = \pi \mathsf{P}$, $\pi \mathbf{1} = \mathbf{1}$ is 1.

5.

A certain calculating machine uses only the digits 0 and 1. It is supposed to transmit one of these digits through several stages. However, at every stage, there is a probability p that the digit that enters this stage will be changed when it leaves and a probability $q = 1 - p$ that it won't. Form a Markov chain to represent the process of transmission by taking as states the digits 0 and 1.
(i) What is the matrix of transition probabilities?
(ii) For $n = 2, 3$ find the probability that the machine, after n stages of transmission produces the digit 0 (i.e., the correct digit).
(iii) Find the stationary distribution(s) of the Markov chain.

Solution. (i) The picture we have here is like this:

$$\xrightarrow{X_0} \boxed{\text{STAGE 1}} \xrightarrow{X_1} \boxed{\text{STAGE 2}} \xrightarrow{X_2} \boxed{\text{STAGE 3}} \xrightarrow{X_3} \cdots$$

where the X_n take values in $S = \{0, 1\}$ and where each stage either flips the entering digit or keeps it the same, independently of the others. So

$$\mathsf{P} = \begin{pmatrix} q & p \\ p & q \end{pmatrix}$$

(ii) The probability that that the machine will produce 0 in two stages if it starts with 0 is $p^2 + q^2$.

(iii) Any stationary distribution $\pi = (\pi(0), \pi(1))$ satisfies

$$\pi(0)q = \pi(1)p, \quad \pi(0) + \pi(1) = 1,$$

whence the stationary distribution is unique with

$$\pi(0) = \pi(1) = 1/2.$$

6.

Mr Baxter's bank account (see Example in lecture notes) evolves, from month to month, according to the rule

$$X_{n+1} = \max(X_n + D_n - S_n, 0), \quad n = 0, 1, 2, \ldots$$

where $X_0, D_0, S_0, D_1, S_1, D_2, S_2, \ldots$ are independent random variables. Suppose that the S_n have the common distribution

$$P(S_n = 0) = p, \quad P(S_n = 1) = q = 1 - p.$$

Suppose that the D_n have the common distribution

$$P(D_n = k) = \alpha^k (1 - \alpha), \quad k = 0, 1, 2, \ldots,$$

where $0 < p, \alpha < 1$.
(i) Compute the one-step transition probabilities
(ii) (Attempt to) draw the transition diagram.
(iii) What is the form of the transition probability matrix?
(iv) Write down the equation that you need to solve in order to compute the stationary distribution (if it exists!) of the Markov chain.

Solution.

(i) If $X_n = x \geq 1$ then X_{n+1} takes values in the set $\{x-1, x, x+1, x+2, \ldots\}$. We have

$$\begin{aligned} p_{x,x-1} &= P(X_{n+1} = x-1 | X_n = x) \\ &= P(D_n - S_n = 1) = P(D_n = 0, S_n = 1) = (1-\alpha)q. \end{aligned}$$

For $m \geq 0$,

$$\begin{aligned} p_{x,x+m} &= P(D_n - S_n = m) \\ &= P(D_n = m, S_n = 0) + P(D_n = m+1, S_n = 1) \\ &= \alpha^m (1-\alpha)p + \alpha^{m+1}(1-\alpha)q. \end{aligned}$$

(It is always good to check that $\sum_y p_{x,y} = p_{x,x-1} + \sum_{m \geq 0} p_{x,x+m} = 1$.)
If $X_n = 0$ then X_{n+1} takes values in the set $\{0, 1, 2, \ldots\}$. For $m \geq 1$, e have

$$p_{0,m} = \alpha^m (1-\alpha)p + \alpha^{m+1}(1-\alpha)q,$$

as before. But

$$\begin{aligned} p_{0,0} &= P(\{D_n = 0, S_n = 0\} \text{ or } \{D_n = 1, S_n = 1\} \text{ or } \{D_n = 0, S_n = 1\}) \\ &= (1-\alpha)p + \alpha(1-\alpha)q + (1-\alpha)q \\ &= (1-\alpha)(1+\alpha q). \end{aligned}$$

(ii)-(iii) In class.

(iv) For all $x \geq 1$ we have that any stationary distribution satisfies

$$\begin{aligned} \pi(x) &= \sum_y \pi(y) p_{y,x} \\ &= \pi(x+1)p_{x+1,x} + \pi(x)p_{x,x} + \sum_{m=1}^{x} \pi(x-m)\pi_{x-m,x}, \end{aligned}$$

one equation for each $x \geq 1$, together with

$$\sum_{x=0}^{\infty} \pi(x) = 1.$$

(We need not consider an equation for $\pi(0)$, because there is always a redundancy amongst the equations $\pi = \pi \mathsf{P}$ owing to, precisely, the last normalisation condition.)

7.

Definition: A Markov chain is called an **actuarial chain** if:
(a) It has a finite number of states (typically less than 10).
(b) It is an absorbing chain.
(c) It has time-varying one-step transition probabilities. [1]

Consider an urn with n balls, out of which m are red and $n-m$ blue. (See exercise of earlier homework.) Assign the value 1 to a red and 0 to a blue ball. Start picking the balls, one by one (without replacement), and let S_t be the sum of the values of the balls you have picked up to the t-th step. Start with $S_0 = 0$ and observe that, obviously, $S_n = m$. Define $S_t = n$ for all $t > n$.
(i) Show that $(S_t, t = 0, 1, \ldots)$ is a Markov chain.
(ii) Compute $P(S_{t+1} = j | S_t = i)$ for all values of states i, j and 'times' t. Observe that the result depends on t.
(iii) Conclude that the chain is an actuarial one.

Solution.

(i)+(ii)+(iii) Recall that S_t denotes the number of red balls picked by time t. So at time t (i.e. right after the t-th selection), we have an urn that contains $n-t$ balls out of which $m - S_t$ are red. Hence, assuming that $S_0 = 0$, we have

$$P(S_{t+1} = x+1 | S_t = x, S_{t-1}, S_{t-2}, \ldots) = P(S_{t+1} = x+1 | S_t = x) = \frac{m-x}{n-t},$$

as long as $m - x \le n - t$, $0 \le x < m$. On the other hand,

$$P(S_{t+1} = x | S_t = x, S_{t-1}, S_{t-2}, \ldots) = P(S_{t+1} = x | S_t = x) = 1 - \frac{m-x}{n-t}.$$

8.

Consider the Ehrenfest chain with N molecules.
(i) Compute its stationary distribution π.
(ii) Suppose that $N = 10^{20}$ molecules. Show that

$$\pi(N/2) \approx 10^{-10} = 0.0000000001$$

but

$$\pi(N/2.0001) \approx 10^{-60,000,000,000}$$

which is the number 0.0.......01 (i.e. $60,000,000,000$ zeros after the decimal point).
Hint: You may use Stirling's approximation, i.e. $n! \approx n^n e^{-n}\sqrt{2\pi n}$ *for large* n.
(iii) If it takes about 1 millimetre to write the symbol 0 show that you need about 5 Earths to write the last number explicitly.
Hint: The diameter of the Earth is about 12 thousand km.
(iv) If you wanted a faithful plot of $\pi(i)$ as a function of i, what shape would the plot have? Sketch it.

[1] Strictly speaking, an actuarial chain is a continuous-time chain, but we shall waive this requirement for the purposes of this part of the course.

Solution. (i) Letting x be the number of molecules in the left room, we have, for $1 \le x \le N-1$,

$$p_{x-1,x} = \frac{N-x+1}{N}$$
$$p_{x,x-1} = \frac{x}{N}$$

Fix a state $1 \le x \le N$ and Partition the state space $S = \{0, 1, \ldots, N\}$ into $A = \{0, 1, \ldots, x-1\}$ and $A^c = \{x, x+1, \ldots, N\}$. If π is a stationary distribution, we have that the flow

$$F(A, A^c) = \sum_{i \in A} \sum_{j \in A^c} \pi(i) p_{i,j} = \pi(x-1) p_{x-1,x} = \pi(x-1) \frac{N-x+1}{N}$$

from A to A^c must be equal to the flow

$$F(A^c, A) = \sum_{j \in A^c} \sum_{i \in A} \pi(j) p_{j,i} = \pi(x) p_{x,x-1} = \pi(x) \frac{x}{N}$$

from A^c to A. Hence

$$\pi(x) = \frac{N-x+1}{x} \pi(x-1),$$

for all $1 \le x \le N$. Therefore $\pi(x-1) = \frac{N-x+2}{x-1}$, $\pi(x-2) = \frac{N-x+3}{x-2}$, etc, and so

$$\begin{aligned}
\pi(x) &= \frac{N-x+1}{x} \frac{N-x+2}{x-1} \cdots \frac{N}{1} \pi(0) \\
&= \frac{N!/(N-x)!}{x!} \pi(0) \\
&= \binom{N}{x} \pi(0).
\end{aligned}$$

But

$$1 = \sum_{x=0}^{N} \pi(x) = \sum_{x=0}^{N} \binom{N}{x} \pi(0) = 2^N \pi(0),$$

whence $\pi(0) = 2^{-N}$ and so

$$\pi(x) = \binom{N}{x} 2^{-N},$$

i.e. the binomial distribution with parameter 1/2.

(ii) If N is large then

$$\begin{aligned}
\binom{N}{x} &\approx \frac{N^N e^{-N} \sqrt{2\pi N}}{x^x e^{-x} \sqrt{2\pi x} (N-x)^{N-x} e^{-(N-x)} \sqrt{2\pi (N-x)}} \\
&= \frac{N^x N^{N-x}}{x^x (N-x)^{N-x}} \frac{e^{-x-N+x}}{e^{-x} e^{-(N-x)}} \sqrt{\frac{2\pi N}{2\pi x 2\pi (N-x)}} \\
&= \left(\frac{N}{x}\right)^x \left(\frac{N}{N-x}\right)^{N-x} \frac{1}{\sqrt{2\pi}} \sqrt{\frac{N}{x(N-x)}}
\end{aligned}$$

If $x = N/2$, we have

$$\binom{N}{N/2} \approx \left(\frac{N}{N/2}\right)^{N/2} \left(\frac{N}{N/2}\right)^{N/2} \frac{1}{\sqrt{2\pi}} \sqrt{\frac{N}{(N/2)(N/2)}}$$
$$= 2^N \frac{1}{\sqrt{2\pi}} \sqrt{\frac{4}{N}} = 2^N \sqrt{\frac{2}{\pi N}}.$$

Therefore

$$\pi(N/2) \approx \sqrt{\frac{2}{\pi N}} = \sqrt{\frac{2}{\pi 10^{20}}} = \sqrt{\frac{2}{\pi}} 10^{-10}.$$

On the other hand, we find, by taking logarithms, that $\pi(N/2.0001)$ is millions of orders of magnitude smaller.

(iii) Easy.

(iv)